Cartographic Map Abstraction Using R Programming
(Literacy Rate, HDI and Poverty Data Interpolation of Nepal)

Yagyanath Rimal[1], Sakuntala Pageni[2]

[1,2]Pokhara University, Tribhuwan, Nepal

E-mail: rimal.yagya@pu.edu.np, pagenisakuntala@gmail.com

---***--

Abstract - *R programming with map data interaction is another research area for data scientists due to high data publication in modern electronic records web applications at the federal administrative structure of Nepal. However the analysis of map and its presentation on web with interactively with data interperation has not much concern with data scientist while website design around the world. The data like central, province, local administrative government bodies many times published data for civil concerned. The dynamic records presentation with local boundaries map having interactive facilities has a new concept using r programming. Here the researcher could easily have developed the local administrative map using GIS shapefile then local level records generally gathered and stored in ms excel will automatically be integrated with this template easily so that local administrative agencies will easily update web site using r programming i.e rpubs. Which discards the registration domain, and knowledge of web application design intelligent site. The best application of this type of data interactive map with data interoperation would be highly applicable to local governance of Nepal where there was a large type of data and records that were developed in sequential order for public concern. The interactive VDC, district, province data commonly highlights data like, education rates and HDI information of any location that could easily be published. The developed model is available (http://rpubs.com/yagyarimal/556607) with interactive website pages quickly utilizing the intelligent markdown with a flash dashboard design template structure.*
*Key Words***:** *Rmarkdown, Human Deveopment Index, Dynamic Web Application.*

1. INTRODUCTION

Web interactive data visualization is one of the most significant parts of data science researchers for the integration of growing data present of local authorities and many more in modern. Maps and figures of information display more than a thousand words. Many times local and central administration institutional records are being poorly displayed on the web without interactive features, therefore, here the researcher tries to developed interactive data presentation using R programming using national education literacy rate and HDI data as sample data sets for demonstrative purpose. The VDC and district Provence region's new authoritative guide of Nepal are being regularly prepared and displayed data. But these data visualization could not have a facility like interactive with map projection. While R is the language for the study of information science 2016, r is the highest information software used 50 percent more than python for information science, there was some reason for its popularity in terms of its free and open source work compared to its vector work where it can work with the entire data analysis facility. Similarly there were huge supporting network community supports more than 30000 packages (Ozgur, 2017). The packages are the function plug and play application for special task supporting units (Paradis, 2005). R studio provides different windows in single application incorporates source window, support window, plots windows and installed information window was exceptionally utilized for information analysis .The latest version 3.6.1 largely supports adaptation running most recent applications like AI, Keras and TensorFlow packages. There were some base bundles and outsider packages were exceptionally accessible. The cran task supports dplyr, tidyr, stringer, httr, ggvis, ggplot2, glossy, rio, and markdown, pacman bundles (Zavadskyy, 2017) free packages for data analysis. Similarly, maps are the representation of points on point from the reference points. The choropleth map use for pointing points

in different regions of a guide for the specific colors and identification. Which shows thickness for each authoritative focuses with a darker or lighter shade. A favorable position is extremely straight forward and shows the spatial dissemination of information very well. A guide investigation is an examination with respect to delineating political maps, military maps, shape lines and so forth., and the one of a kind physical characteristics of maps incorporates scale, title, legend and so on. It is likewise a method for deciphering the message and images of the guide and setting it inside its appropriate spatial and social setting. A guide is a representative delineation stressing connections between components of some space, for example, items, districts, or subjects (DUTTON, 2017). Numerous maps are static, fixed to paper while others are dynamic with data representation. Although information utilizing explanatory analyzing every part of the information given on a map is necessary. Because data are from different sources are accumulated, surveyed, and afterward examined to shape a type of finding or end. Therefore, information processing strategy focus on proper data analysis with appropriate data mining for design publication of organization and institutional data for the purpose of examination, business knowledge and information perceptions for better informatively. A chloroplath map utilizes a uniform subdivision of the world into little standard shapes. Rather, ought to be utilized to outline that are rates, extents, rates, and so cartography maps are intended for information that is non-numeric and has particular classifications.

The human development index (HDI) is a composite measurement of literacy rate, perception income and life expectancy of any country based on United Nation prediction. (Nepal, 2019) The HDI estimates the normal accomplishments in a nation in three essential components of human overall growth, A long life is estimated for future childbirth, as estimated by the grown-up literacy rate and the joined essential, gross enlistment proportion (Feenstra, 2015). The HDI estimation general equation is as HDI= 1/3(life expectancy)+1/3(Education Index)+1/3(GDP) = 1/3 (.76) + 1/3 (.64) + 1/3 (.29) = 0.56. Although, Nepal's HDI index for 2017 is 0.574 which placed the nation in the medium human advancement classification situating it at 149 out of 189 nations and regions. Nepal's HDI index expanded from 0.378 to 0.574, an expansion of 51.9 percent (UNDP, 2018). According to WHO figures, the future in 2017 arrived at 67.7 and 70.8 separately and the present future for Nepal in 2019 is 70.60 years. Nepal has a GII estimation of 0.480, positioning it 118 out of 160 nations in the 2017 list. In 2019, the GDP in Nepal was at around 29.81 billion U.S. dollars. Nepal slipped five spots from 144 to 149 among 189 nations in the Human Development Index positioning in 2017, as indicated by the Human while Afghanistan has been put in the 'low human advancement' gathering. Thus, an educated individual is characterized as an individual (above 6 years or more) who can peruse and compose and perform basic scientific figurings autonomously. In 2001, the national education rate was 54.1%; it differed area insightful from 27.1% (Humla) to 77.2% (Kathmandu). In any case, there was another procedure of computation HDI as well. The measure of products created inside Nepal is amazingly insignificant. Nepal has gone into an administrative political structure with 7 common governments and 753 nearby regions in 2015 (Acharya, 2018). However province-level there were highly unsimilarity is seen as remarkable as far as effective asset use and information, and whenever upheld with more assets has a high possibility of turning into the good example region in Nepal (Government, 2017).

The topical maps in GIS are significant in to choropleth map utilizes detailing zones, for example, districts or statistics points to show information like livelihoods, percent female, or paces of mortality the limits of the zones are wide range of sets of information temperature, weight, precipitation with thickness of color (Dodge, 2008). Information show maps give helpful methods for showing data in a significant manner by the expense of making and printing a guide is high, so its substance is frequently a trade-off between various other points (Tremblay, 2017). Information examination of maps are utilized in investigation for distinguishing proof of growth at the

connection between two disseminations utilizing straightforward overlays. Although cartography is lower cost for basic maps, quicker generation with more prominent adaptability in yield simple scale or projection change, maps can be customized to client needs different uses for computerized information. GIS computer cartography has an important objective of developing map structures that have propelled map format tools, naming location, huge image and text type libraries, interfaces for costly, excellent yield gadgets, despite the fact that it is nothing but an expository method, as opposed to data that GIS considers fast access for traditional as well as creative. Map purposes causes problems when used as hotspots for advanced information most GIS fail to evaluate contrasts (Nation, 2009) between data sets obtained from maps on different scales quirks speculation systems in maps to the information obtained from them such blunders become regularly evident only during the later preparation of computerized information obtained from them (Rhind, 1988). Information files this capacity can be performed much better by a decent GIS because of the capacity to give various and productive cross referencing and looking. Information show instruments electronic showcase offers critical focal points over the paper map capacity concealing free of the imperatives of the printing procedure, capacity to change hues as required for elucidation standout, particular reason items are conceivable and cheap (Tomlinson, 1988). Similarly poverty may be refer as not having enough money to meet basic needs, including food, clothing, and housing. Poverty has no access to school and cannot read. Poverty does not have a job, it is fear for the future, living one day at a time (Narayan, 1999). According to (Nepal N. , 2017) The poverty rate in Nepal has fallen to 21% in 2018. This indicates the success of the government of Nepal so far and its continued efforts for sustainable development. However, most of its population lives below the poverty line.

Therefore here researcher tries to present the map plotting and interactively using r markdown packages for data interoperation using r programming so that data presentation on web with easiest without having knowledge of HTML,CSS, Java Script and PHP for local level administrative of Nepal.

2. INFORMATION ANALYSIS USING R PROGRAMMING

This exploratory review paper attempts to created interactive web site with the design for the presentation of data like HDI, literacy indexes, poverty of 77 districts records with projecting colorful maps after afterhaving loading of GIS Shapefile and excel data records with HDI and literacy rate and Poverty of Nepal. This paper presents p way of plotting various maps of nation and other level informational indexes organizations for open on the web could be effectively configuration site without the knowledge of HTML and other web intelligent applications php, java scripts.net and many more. The last final design of web development can be accessible on http://rpubs.com/yagyarimal/556607. Information analysis of GIS shapefile for map production and interpolation of data like HDI, literacy and poverty index with interactive visualization could be easily developed. This paper tries to overcome conventional way of data presentation on the web of each local unit that could easily be solved and published with spatial investigation are getting accessible. Here are some steps are recommended for doing so. After loading system library packages in r studio.

setwd("C:/Users/user/Desktop/publication/HDImap/NPL")
nepal.adm3.shp<-readOGR(dsn='C:/Users/user/Desktop/publication/HDImap/NPL/NPL_adm4.SHP', layer="NPL_adm4", stringsAsFactors = FALSE)

After environment setting, working directory is setout then loading shape file of district of nepal
nepal.adm3.shp.df<-fortify(nepal.adm3.shp, region = "NAME_4")

The shapefile containing feature geometry and feature attribute data and geom polygons are included and stored in new variable.The spatial manipulation and re-projection are performed on as object polygons. At this moment ggplot2 will render the polygons using geom_polygon expects a standard data frame containing polygon verticies and attribute data were ploted easiy. When each corresponding region shape values were copied from geometric shape plot.

The map object is created with ggplot(data,Nepal.admin3.shp.df,aes(x=longitude,y=latitude having groyp by group then geom_path function plots its line maps.
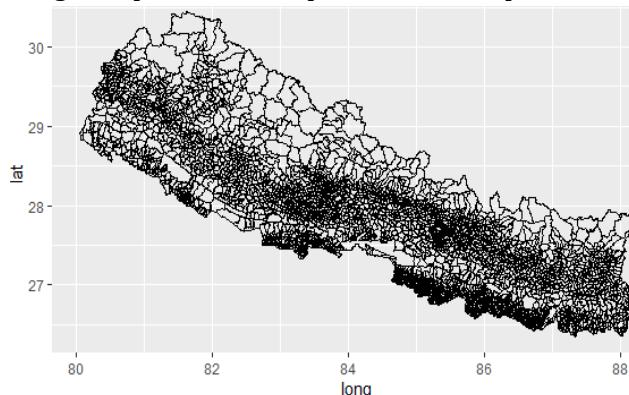

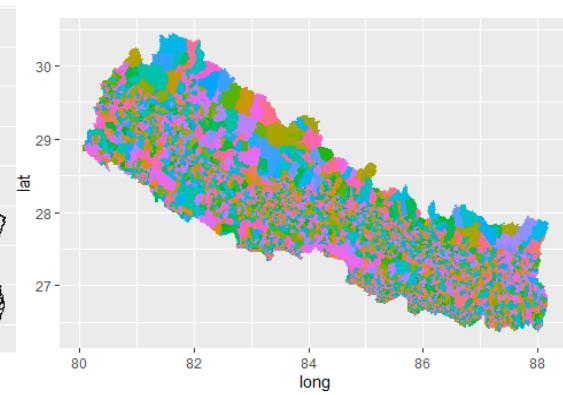Fig 1: VDC Point Map   Fig2: VDC Color Map of Nepal

The VDC color map is easily generated thei filling its id i.e name with corrdinate fixed.

After loading shape file r programming fortify with name_4 or DIST_EN could easily access each chloroplast longitude and latitude points of map are plotted with ggplot with aesthetic longitude and latitude then gemo_ path function plots skeleton ponts. And polygon maps with color code were plotted on the basic of districts or VDC wise records.

From the map data we can easly access its id name and write to our file with row.name is equal false . The %>% pipe symbols is use for making distinct data could easily write in own directory with dropping out table in the working directory. At this point we could easily match the maps points with field excel records.

hpi.data <- read.csv("C:/Users/user/Desktop/publication/HDImap/NPL/districts.csv")

colnames(hpi.data) <- c("id","HDI","Lit")

This lines loads the excel records of district name Hdi and literacy rate of each districts.

>nepal.adm3.shp.df<-merge(nepal.adm3.shp.df,hpi.data, by ="id")

The excel data having human development index and literacy are merged with same file by its id as primary key for data integration as below. So that records could easily ploted with shape of each location.
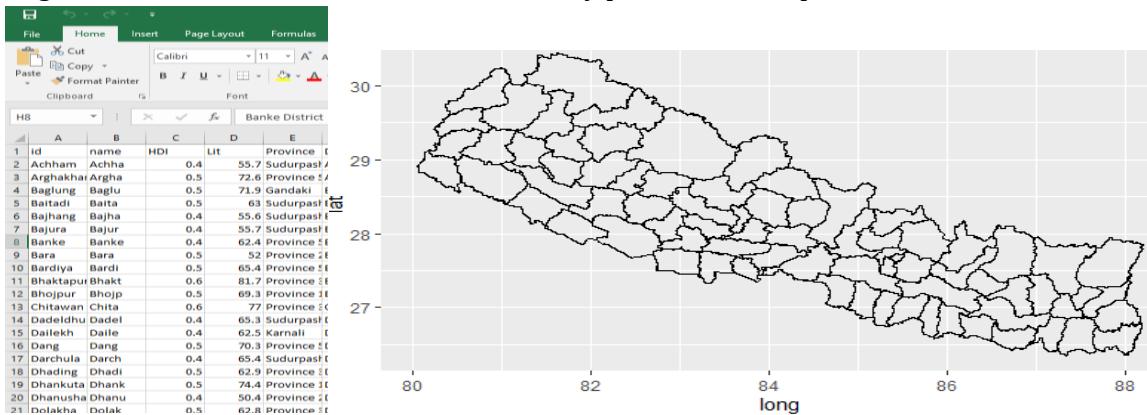

Fig3: 77 District Map of Nepal

The map of 77 districts could be easily developed with ggplot with aesthetic x equal to latitude and y is equal to latitude and grouped by its id i.e district name. when map object is added with geom_path() function formed color chloroplath map.
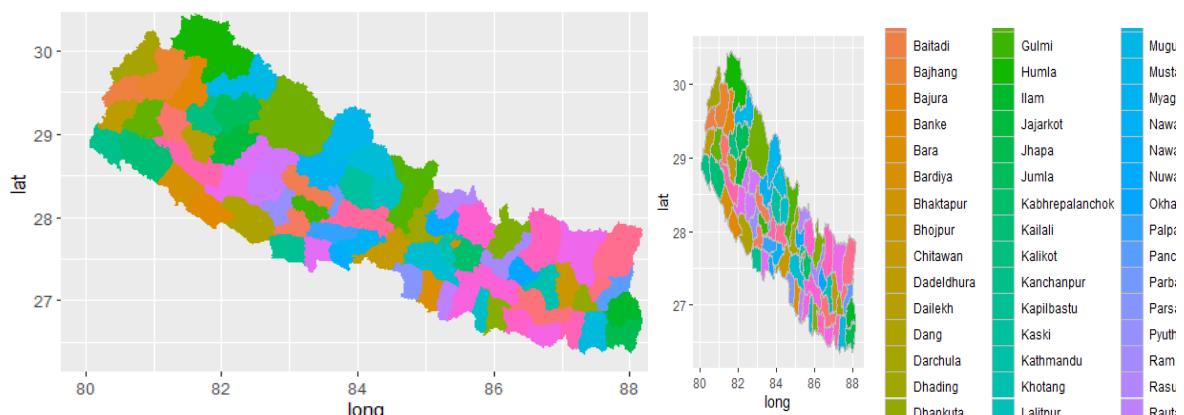
Fig 4: District Color Map with Its Name

The district map with its color legent could easily plot with interacting map objectgeom polygons with filling gradiant color with color bar. After setting low and high color range of different gradient, the geom polygons map plots will easily designed.

The chloroplath map with filling its HDI data sets formed color spead by its hdi weight of 77 district with legend indicating its color bar inside or outside ploted area.
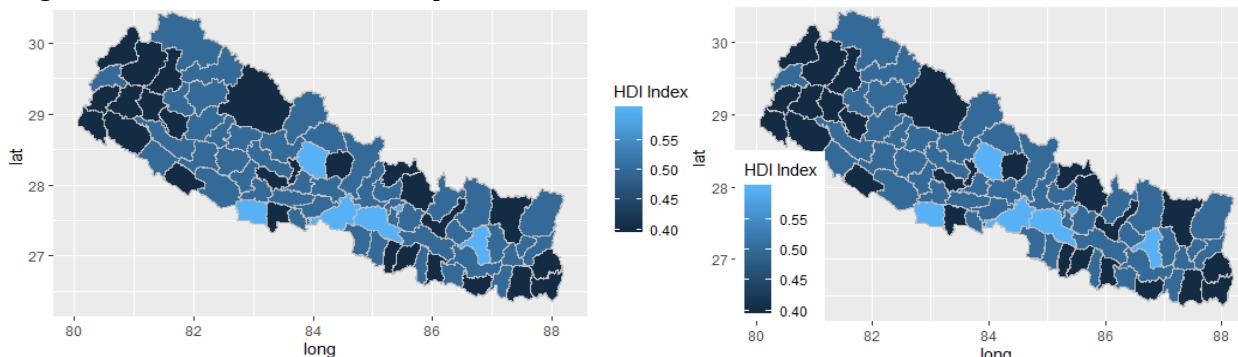

Fig5: HDI Map with gradiant colour with index

After plotting different id the HDI data filled for the geom polygon map of choropleth produces the map legend. The above code produces color bar maps with legend inside or outside of district with HDI data values interoperation with its colors. The darker images indicate with low and light colors with highest HDI index with specialized legend design at side of lap plot.
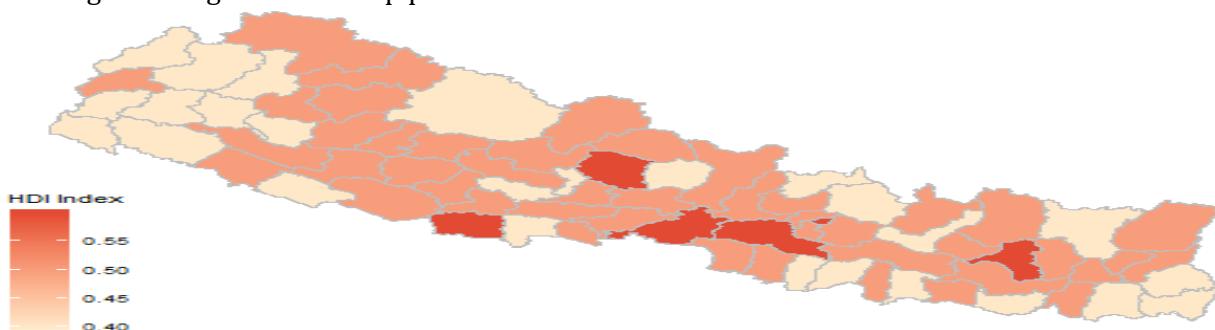

Fig6: Choloropath Map with HDI Indexed

The choropleth map could easily change its color with HDI data using polygon and color change.

centroids <- setNames(do.call(rbind.data.frame, by(nepal.adm3.shp.df, nepal.adm3.shp.df$group, function(x) {Polygon(x[c('long', 'lat')])@labpt})), c('long', 'lat'))

centroids$label <-nepal.adm3.shp.df$id[match(rownames(centroids), nepal.adm3.shp.df$group)]

centroids$label=hpi.data$name

the centroid mf map would calculated using polygon function then its name with five character are set out its label. When map with new centroid data formed with district name.
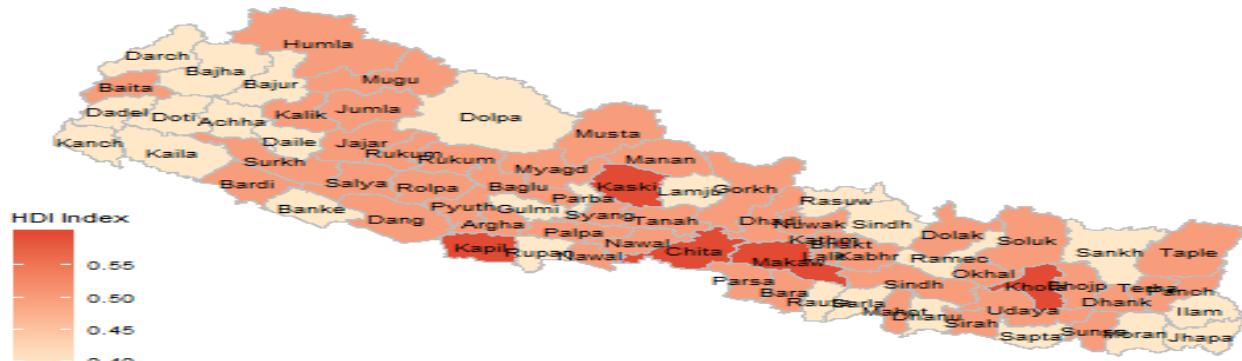

Fig 7: HDI map with its name

The centroid of a map is by using polygon function and filled with its district name. The second map demonstrates only those districts whose human development index is greater thean 0.40 only.

centroids.selected <- centroids[centroids$label %in% (hpi.data[hpi.data$HDI>0.4,]$id),]

When plot with map formed demonstrate those district with light and darker color map with only those district whose having more than 40 percent HDI value in their exvel data.
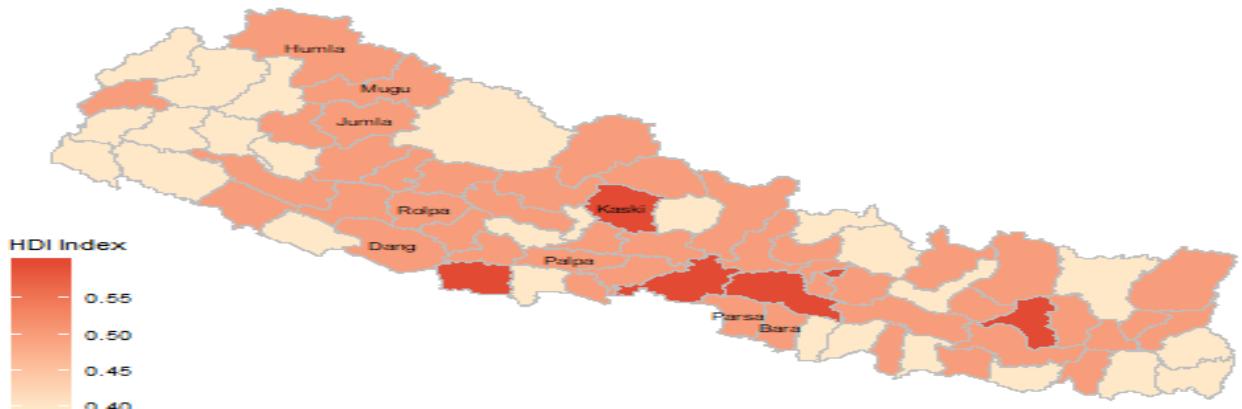

Fig7: HDI with more than 40 HDI

From the above map district like kaski kapilbastu, chitwan, makanwanpur and khotang has highest human index district similarly district like dolpa, lunging, dolkha, solukhumbu, banke , doti, bajura, kailali kantanchanpur and dailekh has lowest human development index districts of Nepal.

Similarly literacy data interaction with map object formed Literacy map as below.
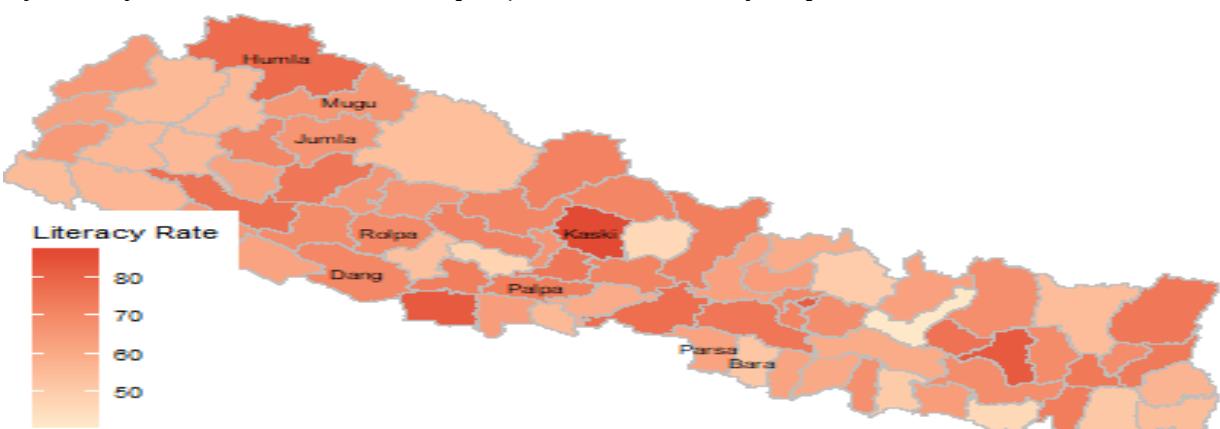

Fig9: Literacy Map ofNepal

From the above map district like kaski, kapilbastu,humla has highest literacy rate where as lamgung ramachhap, pyuthan and gulmi has lowest literacy district in Nepal.

Similarly when map object with poverty excel data feeds plots the map of poverty index.
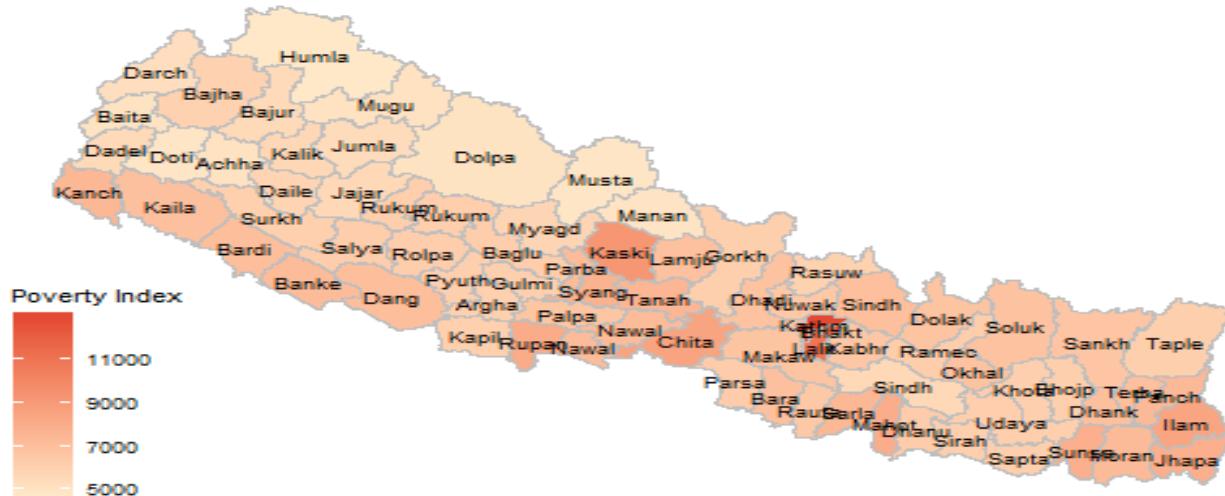

Fig 10: Poverty map of Nepal

This poverty map could be easily plots with displaying the highest color for lowest poverty indicitator of 77 districts of Nepal The districts namely humla,doti achham, dolpa mustang,manang and sindhuli has higest poverty ratched districts of Nepal.

3. CONCLUSION

This exploratory data interoperation article tries the first time to integrate geographical maps interactive web applications for presenting various data in modern for each location. This is steps to integrates District Human Development Index and Literacy Rate of Nepal will be easily plotted using a chloroplast map of shapefile and excel data table. The chloroplast map uses longitude and latitude data to represent each administrative location of Nepal where there was much data information may project district, village and locality, therefore, CBS like government offices will highly be recommended for using this concept to present data information using r programming. This output of this integrated work is available http://rpubs.com/yagyarimal/556607 link for the demonstration between map and data interpretation.

REFERENCES

[1]   Acharya, K. K. (2018). Local Governance Restructuring in Nepal: From. Dhaulagiri Journal of Sociology and Anthropology Vol. 12, 2018 PP 37.

[2]   Chetri, T. B. (2017). Federal Democratic Republic Nepal:.

[3]   Dodge, M. (2008). Understanding Cyberspace Cartographies:. A thesis submitted for the degree of Doctor of Philosophy.

[4]   DUTTON, R. W. (2017). Generalising spatial data and dealing with multiple. 125.

[5]   Feenstra, R. C. (2015). The Next Generation of the Penn World Table. University of California, Davis and NBER, 3.

[6]   Gajdoš, V. U. (2019). Hierarchical Hexagonal Clustering and Indexing. Research Gate.

[7]   Government, N. A. (2017). Diagnostic Study of Local Governance in The Asia Foundation. Diagnostic Study of Local Governance in Federal Nepal was implemented with support.

[8]   Kennelly, J. S. (2010). Illuminated Choropleth Maps. Annals of the American Association of Geographers 100(3):513-534, Reserch Gate.

[9]   Narayan, D. (1999). Can Anyone Hear Us ?Voices From 47 Countries. Poverty Group, PREM World Bank, December 1999.

[10] Nation, U. (2009). HANDBOOK ON GEOSPATIAL INFRASTRUCTURE .

[11] Nepal, N. (2017). DraftNepal's Sustainable Development Goals Baseline ReportJune2017. National Planning Commission,Government of Nepal.

[12] Nepal, U. (2019). Human Development Index and its components. Human Development Report.

[13] Ozgur, C. (2017). MatLab vs. Python vs. R. Journal of data science: JDS 15(3):355-372.

[14] Paradis, E. (2005). R for Beginners. Emmanuel Paradis (12th September 2005).

[15] Rhind, D. (1988). Personality as a factor in the development of a discipline: the example of computer- assisted cartography," American Cartographer 15:277-90. . Examines the history of the digital revolution in cartography and the effect of key personalit.

[16] Tomlinson, R. (1988). The impact of the transition from analogue to digital cartographic representation. American Cartographer 15:249-62. An overview from a pioneer of GIS. .

[17] Tremblay, K. (2017). ASSESSMENT OF HIGHER EDUCATION. OECD.

[18] UNDP. (2018). Human Development Indices and Indicators: 2018 Statistical Update. 1.

[19] Zavadskyy, V. (2017). The goal of ggvis is to make it easy to describe interactive web graphics in R. https://gist.github.com/VladislavZavadskyy/e31ab07b03a5c22b11982c49669a400b.